Skill assessment of sub-seasonal forecasts of temperature and precipitation

Andrea Manrique-Suñén[1], Nube González-Reviriego[1], Nicola Cortesi[1], Raül Marcos[1], Albert Soret[1], Francisco J. Doblas-Reyes [1,2]
(1) Barcelona Supercomputing Center (BSC), Barcelona, Spain; (2) Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

1. Introduction

Sub-seasonal climate predictions provide information in the time range from 2 weeks to several weeks into the future. For hydrological applications, such as water management for hydropower or irrigation in agriculture, this time range is very valuable for decision making. A forecast quality assessment of temperature and precipitation has been performed by comparing a set of predictions in the past (hindcasts) of 2 sub-seasonal systems with a reference.

2. Methodology

Sub-seasonal systems

Sub-seasonal prediction systems vary greatly in their configuration of both the forecast and the hindcast. A summary of the main characteristics of the two predictions systems analysed is shown in Table 1. Data were obtained from the Subseasonal to Seasonal Project Database (Vitart et al. 2017), and are previously interpolated to a common grid of 1.5° (240x121).

System	Version	Reference	Forecast		Hindcast			
			Frequency	Ensemble size	Frequency	Ensemble size	Type	Years
ECMWF	CY41R2	Vitart (2004)	Mon/Thu	51	Mon/Thu	11	On the fly	20 previous years (For this study 1996-2015)
NCEP	CFSv2	Saha et al. (2014)	Daily	16	Daily	4	Fixed	1999-2010

Table 1. Characteristics of the two sub-seasonal systems

Reference datasets

- **2m T**: ERA-Interim reanalysis (Dee et al. 2011)
- **Precipitation**: Multi-Source Weighted-Ensemble Precipitation dataset (MSWEP) (Beck et al. 2017)

Products

Due to the differences in the setup of the systems, the methodologies applied to compute the anomalies and the verification measures are different for ECMWF and NCEP systems. This means that the skill scores cannot be directly compared. Weekly averages were calculated from day 5 onwards, producing 4 forecast times: **week 1** (days 5-11), **week 2** (days 12-18), **week 3** (days 19-25) and **week 4** (days 26-32). **Anomalies** of the weekly averages were computed for each start date and forecast time in the hindcast and analogously in the reference dataset. In the case of NCEP, due to the small number of hindcast years (12) and members (4), forecasts from the day before and the day after were used to compute the anomaly. A timeline for a sample month (January) with the corresponding integrations and start dates issued by each model is shown in Figure 1. Verification measures, correlation of the ensemble mean (EnsCorr) and Fair Continuous Ranked Probability Skill Score (FairCRPSS) (Ferro, 2014) were calculated for each system, for each forecast time and each month by combining together all start dates within each month.

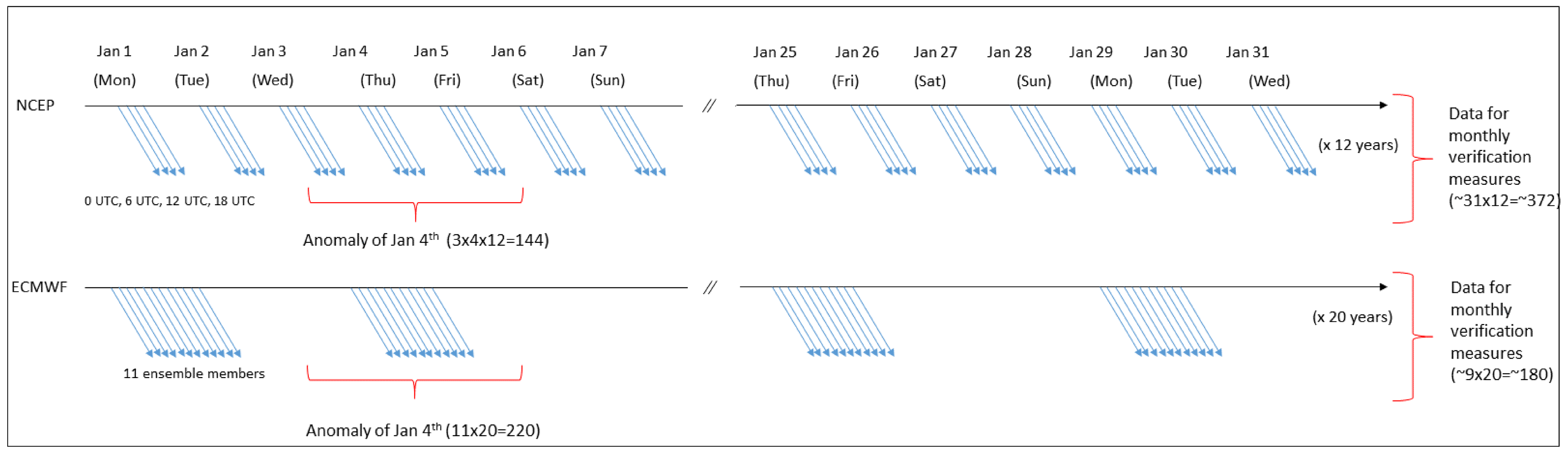

Figure 1. Schematic of the hindcast configuration of each system for one month (January as an example). The blue arrows indicate the model integrations on each start date

3. Results

Verification measures are shown for January and July over Europe. Positive coefficient of correlation of the ensemble mean indicate linear correlation of predictions with the reference. FairCRPSS ranges from -∞ to 1, positive values indicate that the model´s prediction has higher skill than a forecast based on climatology.

Acknowledgements
The research leading to these results has received funding from the EU H2020 Framework Programme under grant agreement n° H2020 GA 641811
Hindcasts of sub-seasonal predictions were obtained from the Subseasonal to Seasonal (S2S) Project database

Bibliography

Beck, Hylke E., Albert I. J. M. van Dijk, et al. 2017. "MSWEP: 3-Hourly 0.25° Global Gridded Precipitation (1979-2015) by Merging Gauge, Satellite, and Reanalysis Data." *Hydrology and Earth System Sciences* 21(1):589–615.

Dee, D. P. et al. 2011. "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System." *Quarterly Journal of the Royal Meteorological Society* 137(656):553–97.

Ferro, C. A. T. 2014. "Fair Scores for Ensemble Forecasts." *Quarterly Journal of the Royal Meteorological Society* 140(683):1917–23. Retrieved July 6, 2017 (http://doi.wiley.com/10.1002/qj.2270).

Vitart, F. 2004. "Monthly Forecasting at ECMWF." *Mon. Wea. Rev.* 132:2761–79.

Vitart, F. et al. 2017. "The Subseasonal to Seasonal (S2S) Prediction Project Database." *Bulletin of the American Meteorological Society*.

Saha, Suranjana et al. 2014. "The NCEP Climate Forecast System Version 2." *Journal of Climate* 27(6):2185–2208.

2m temperature

Both verification measures show good skill for week 1, but as lead time increases skill deteriorates. Generally, higher skill is found over the ocean (especially ECMWF). Over the continent in winter there are some areas (Central-Eastern Europe and Russia) where ECMWF presents high correlation coefficients up to week 4. In summer, correlation values are positive in the Mediterranean area (week 4 in ECMWF and week 3 NCEP).

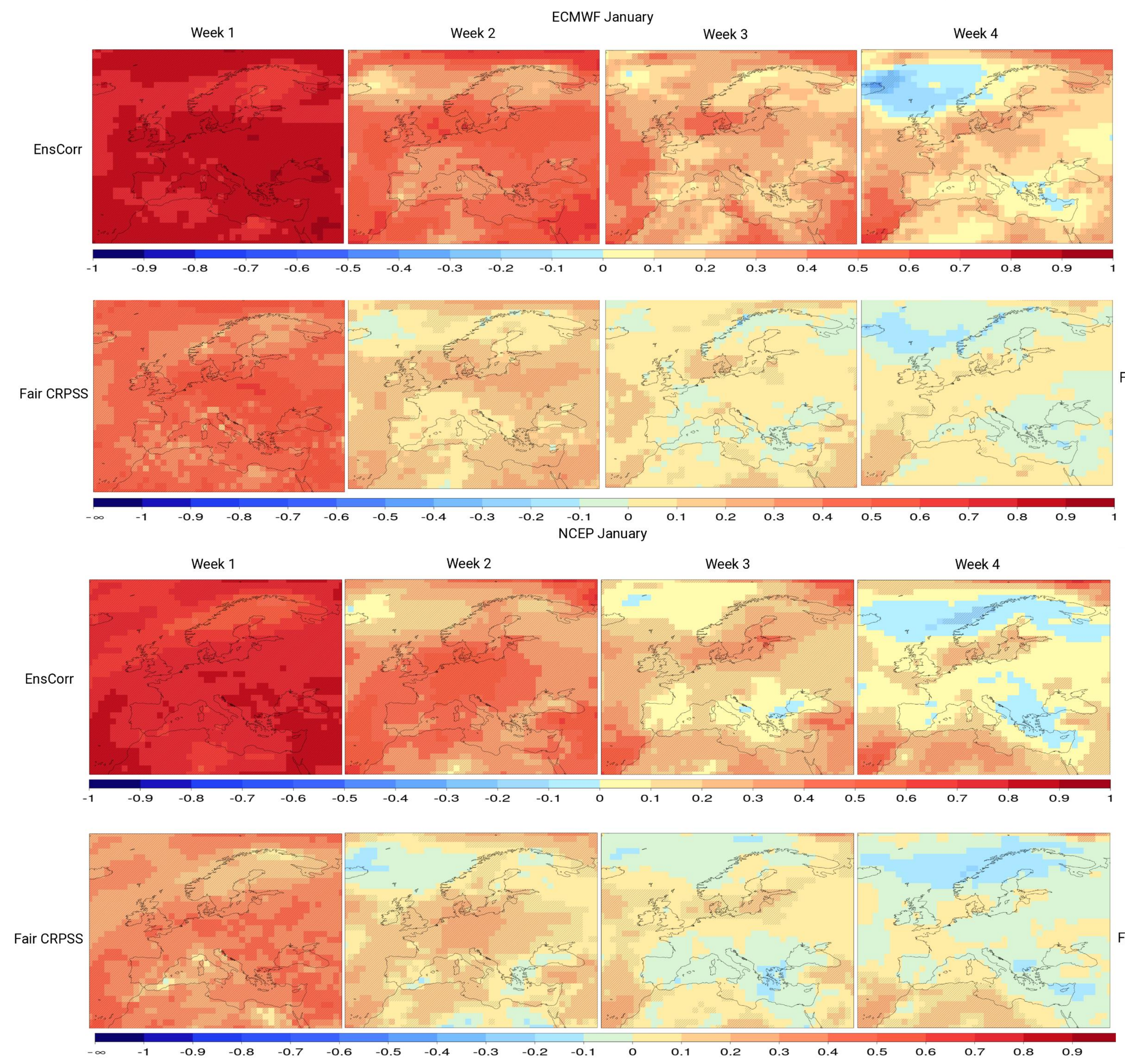

Figure 2. Verification measures for 2m T for ECMWF (top) and NCEP (bottom) for January.

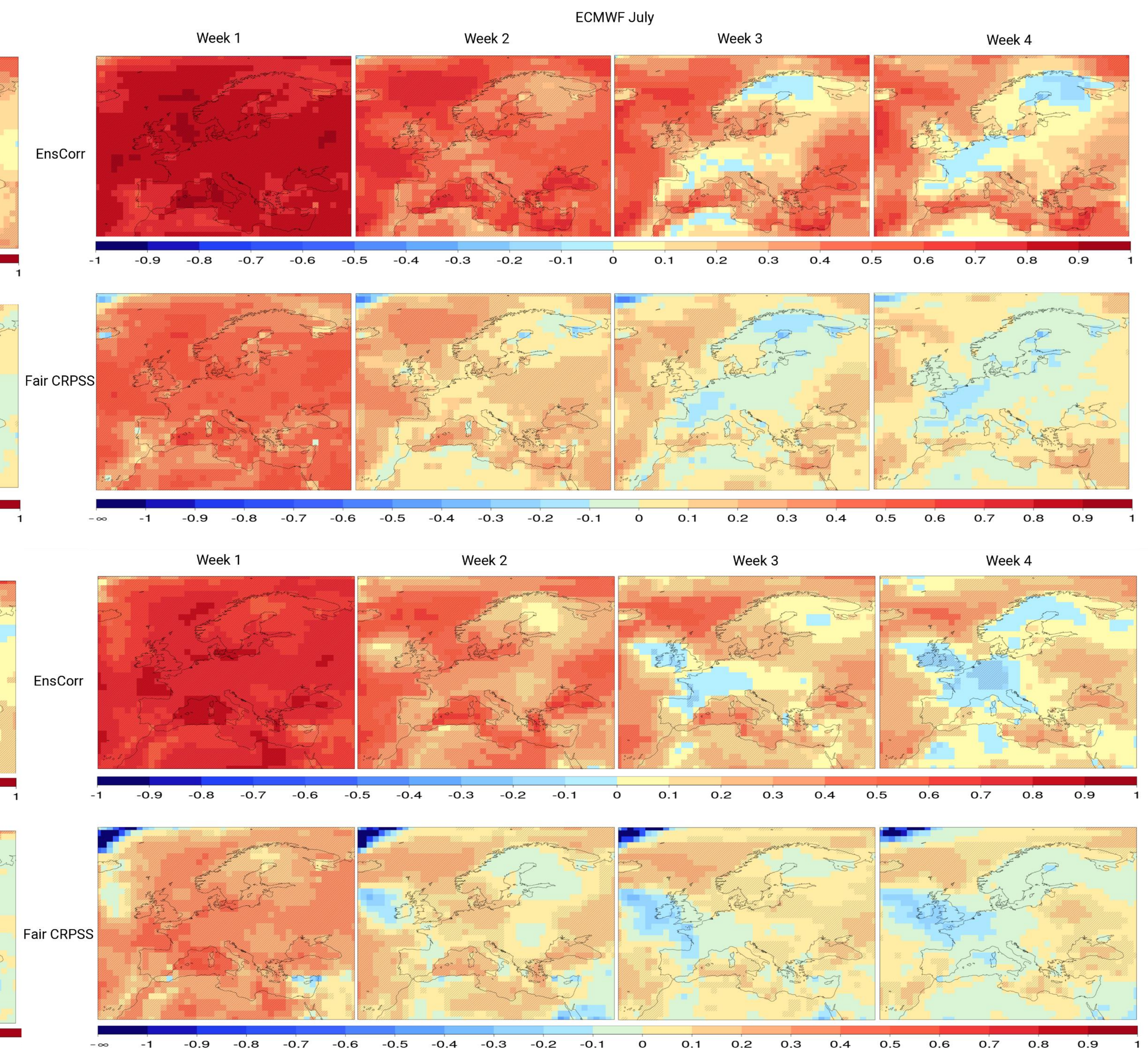

Figure 3. Verification measures for 2m T for ECMWF (top) and NCEP (bottom) for July.

Precipitation

Precipitation predictions show less skill than those of temperature and the verification measures show a noisier distribution. In the case of NCEP, only EnsCorr is shown as FairCRPSS showed negative values.

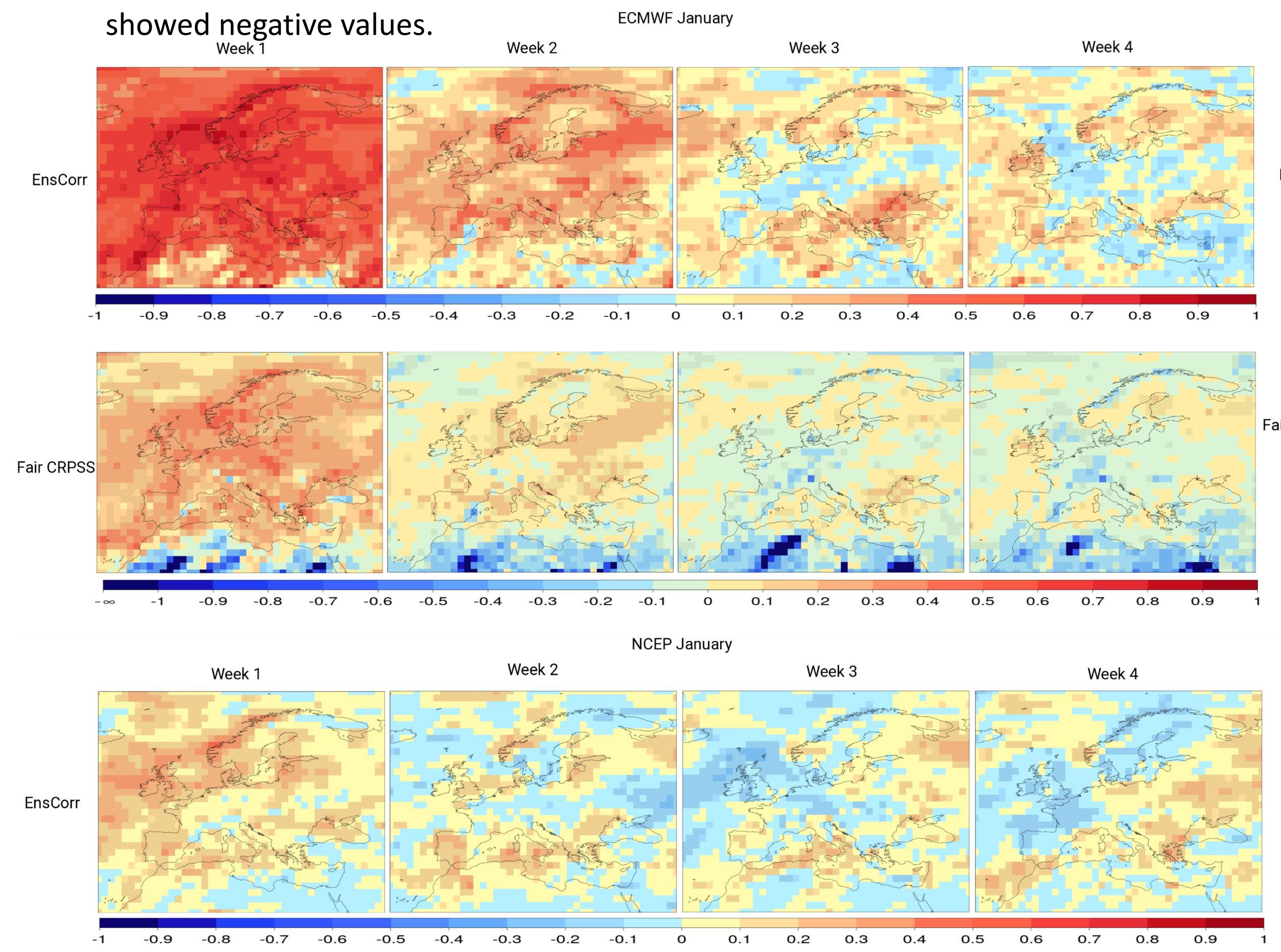

Figure 4. Verification measures for precipitation for ECMWF (top) and NCEP (bottom) for January.

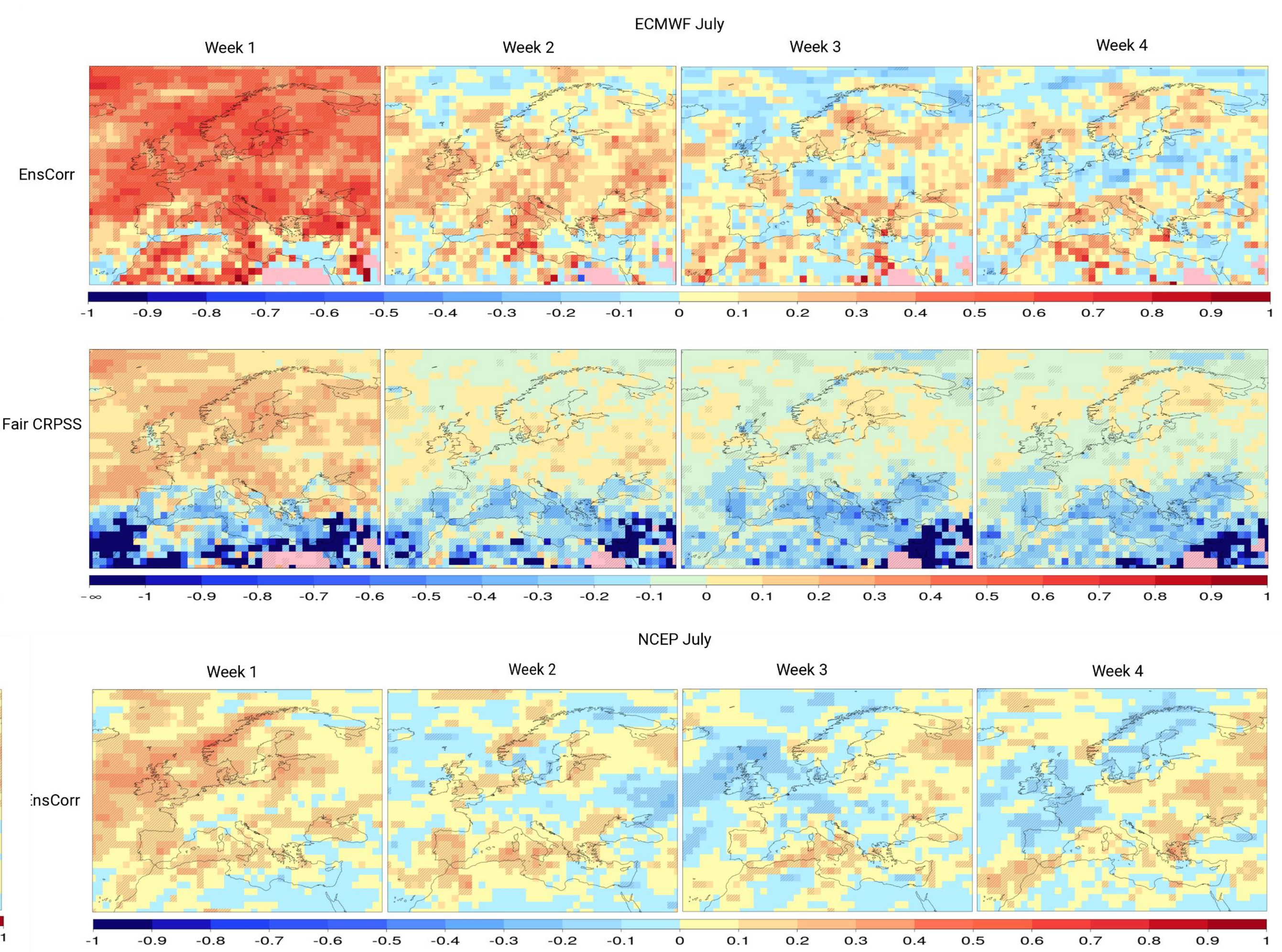

Figure 5. Verification measures for precipitation for ECMWF (top) and NCEP (bottom) for July. Pink colour indicates areas where the climatological value is zero and it is not possible to compute the verification measures.

4. Conclusions

Temperature predictions showed significant skill in continental areas up to day 18. Some specific areas (Central-Eastern Europe) in winter and around the Mediterranean Sea in summer presented positive values of correlation of the ensemble mean for longer lead times, suggesting windows of opportunity in these areas. Precipitation forecasts showed low skill after day 11, some calibration techniques will be explored.

5. Future work

- Bias correction will be applied to the raw predictions
- The forecast quality assessment will be conducted for 10 m wind in the context of S2S4E and NEWA projects.